THE SAFETY IMPACT OF SUBSTITUTING MANUAL OPERATION OF IN-VEHICLE TASKS WITH VOICE CONTROL

Arihiro Isomura

Mitsubishi Motors Corporation
Japan
Paper No.84

ABSTRACT

In respect of manual control car navigation system, control safety of in-vehicle device is assured by restricting the driver's complex operations such as setting or revising destination by operating cursor switch while driving. Because the driver's workload when using voice-controlled system is small, voice operation is adopted as an effective operation method to reduce the driver distraction. In this research paper, voice control, a method of reducing driver’s distraction, is investigated and researched by testing under different driving and control conditions to determine the influence of the loads.

INTRODUCTION

The number of car navigation systems in Japan continues to grow each year. The total number of car navigation systems shipped to Japan exceeded 10 million in 2002. Car navigation systems are becoming standard equipment in the Japanese and European automobile industry. Most car navigation systems now manufactured for Europe and Japan with advanced functions are equipped with voice control systems.

The manual control system promotes safety by restricting the driver's complex operations such as setting or revising destination by operating cursor switch while driving [1]. However, safety is not ensured when using a voice-controlled system, because there is no restriction on the number of control functions the driver can request. The increase in the driver’s workload when using a voice-controlled system is minimal, but not inconsequential.

During testing, the driver travels a simulated curved course while using a voice control system. The road curve rate for measuring driving workload and the number of voice commands are controlled.

TEST

Device Outline

As shown in Figures 1 and 2, Flat belt driving simulator is used to measure safety and to avoid unnecessary influence from another vehicle. Main features of this simulator are as follows;

By using an actual vehicle, steering reaction and engine noise can be measured.

The simulation is completed using a 100-degree cylinder shaped screen. (See figure 3)

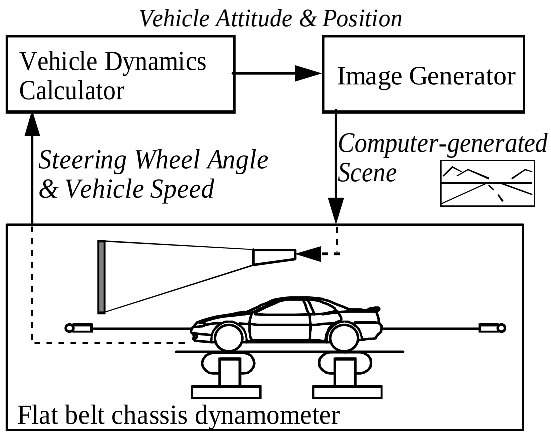


Figure 1. Driving simulator block diagram.

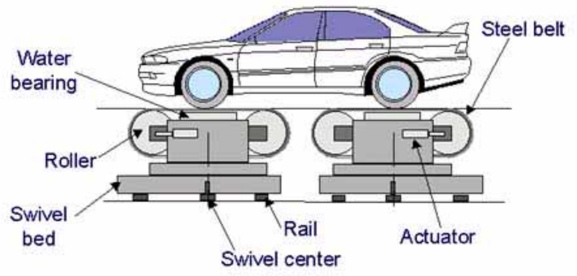


Figure 2. Flat belt chassis dynamometer.

Figure 3. Picture of simulation test.

The navigation system used for this testing is a standard voice-controlled system with a remote-controlled switch.

Figure 4 shows where the navigation system is installed inside the vehicle. A push-to-talk switch for voice control is installed in the steering wheel, and a microphone is installed in the sun visor.

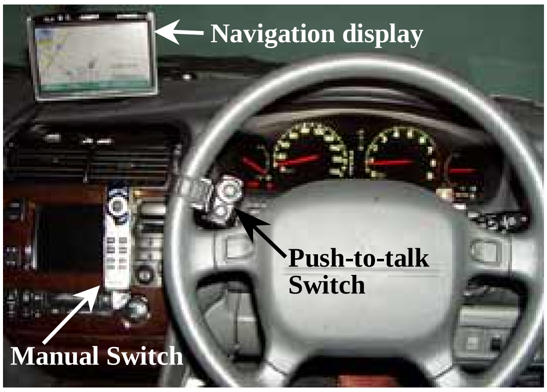


Figure 4. Picture of installed navigation system.

The procedure for using the voice control is as follows:

1. Press push-to-talk switch.
2. State command.
3. Navigation system recognizes command and pronounces recognized command by voice synthesis.
4. Operation is performed.

The switch for the manual control is installed where it can be most easily reached near the radio controls. (Figure 4)

Test Method

Testing is performed based on the dual task method in which the simulator test is the main task and the car navigation control is the sub-task. The test result is evaluated based on vehicle performance and driver performance.

The sub-tasks are shown in Table 1. Limiting the number of voice commands to two, four, or six controls the data conclusions drawn from the testing of the voice command system. In order to compare manual control and voice control systems, voice control sub-task 2 and manual control sub-task 2 are developed. The workload of the main task is controlled by four different curve rates as shown in Figure 5.

<u>Test Procedure</u>

The driving course is composed of a straight road and a curved road. The driver controls the starting time of the sub-task when navigating the curved road. Drivers are asked to increase driving speeds up to 100 km/h on the straight road and to keep driving speed at 100 km/h on the curved road. In order to evaluate the influence of the sub-task on the driver, normal driving tests without the navigation system are practiced as well.

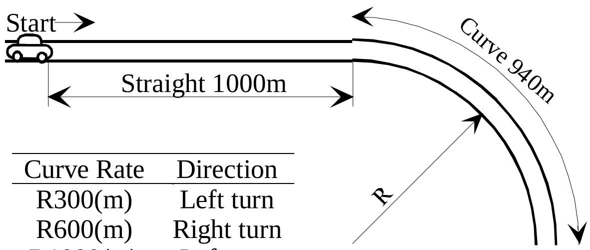
Drivers are allowed to practice simulator driving and navigation control before the test to ensure they are familiar with the operation.

If the navigation system does not recognize the driver's command, the test is started over from the beginning.

Subject drivers are four males ranging in age from early 20s to late 30s.

Table 1.
Sub-task

Voice Control Sub-task	Number of command	Detail of Command
Sub-task 1	2	1 [Wide area] 2 [Wide area]
Sub-task 2	4	1 [Menu] 2 [VICS] 3 [NHK VICS] 4 [Character information]
Sub-task 3	6	1 [Menu] 2 [Destination] 3 [Search of Destination] 4 [Address] 5 [Kanagawa Prefecture] 6 [Yokohama City]
Manual Control Sub-task	Number of Control	Context of Control
Manual Control Sub-task 2	4	Press [menu] button once and [action] button 3 times. (There is not operation with joystick)



Curve Rate	Direction
R300(m)	Left turn
R600(m)	Right turn
R1000(m)	Left turn
R2000(m)	Right turn

Figure 5. Diagram of driving course.

Measurement Item and Analysis Method

Behavior of Subjects

Sub-task completion times of voice control is measured from the moment a driver pushes the switch to the moment the final screen of the sub-task is displayed on the navigation screen.

Sub-task completion times of manual control is measured from the moment driver's hand leaves the steering wheel to the moment the final screen of the sub-task is displayed on the navigation screen.

Subjective Evaluation

Anxiety levels during control testing was measured in seven levels:

7: Extremely nervous
6: Very nervous
5: Nervous
4: Neither
3: Comfortable
2: Very comfortable
1: Extremely comfortable

Mental workload is evaluated based on the Japanese version of the NASA Task Load Index (NASA-TLX)[2]. In addition to the NASA-TLX, a supplementary scale entitled, "overall workload" is applied whereby the subject rates his stress level overall.

Vehicle Behavior

The standard deviation of the steering angle and the range of lateral positioning of the vehicle in the driving lane are both measured during the sub-task testing. Data is sampled every 50 milliseconds.

THE TEST RESULTS

The test results are evaluated based on the evaluation indices for curve rate and sub-task influence.

Behavior of Subjects

Sub-Task Completion Time

Task completion time increases as the voice control command number increases.

When comparing the same task completed via voice control versus manual control (sub-task 2), voice control tasks take longer due to the length of time the driver speaks command.

Task completion time is not influenced by the curve rate under the current curve rate range. This result is measured when driving becomes stable. It also explains that navigation control was not interrupted by driving.

The data section without sub-task is defined as follows. Starting time of data is eight seconds after entering curve area. This is the same as the average starting time of the sub-task. The data length is 23 seconds, the same as the sub-task completion time for sub-task 3.

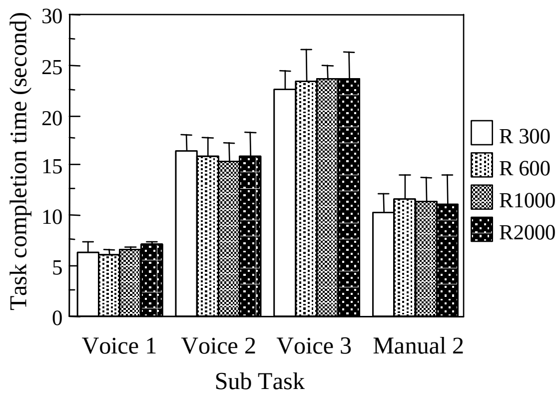


Figure 6. Task completion time.
(Graph shows average of all subjects. Error bar shows standard deviation)

Table 2.
ANOVA (task completion time)

Source	DF	Sum-of-squares	Mean-square	F-Ratio	p-value
Curve(A)	3	3.143	1.048	0.262	0.8526
Sub-task(B)	3	2471.386	823.795	205.821	<0.0001
A*B	9	9.370	1.041	0.260	0.9822
Error	48	192.119	4.002		

Difference between means		p-value
Voice 1 - Voice 2	-9.371	<0.0001
Voice 1 - Voice 3	-16.779	<0.0001
Voice 1 - Manual 2	-4.546	<0.0001
Voice 2 - Voice 3	-7.408	<0.0001
Voice 2 - Manual 2	4.825	<0.0001
Voice 3 - Manual 2	12.233	<0.0001

Subjective Evaluation

Anxiety during Control

Drivers noted no anxiety during use of the voice control system, except during curve rate R300 of Sub-task 3. This task constitutes the heaviest workloads for both the main-task and the sub-task. Anxiety was measurable on manually controlled sub-task 2 regardless of curve rate.

This result proves navigation control systems can reduce driver anxiety when utilizing the voice as the control method versus manual control. However, as the number of voice commands increases, anxiety can also increase. If the number of voice commands exceeds the number set in this test, an increase in anxiety is measured.

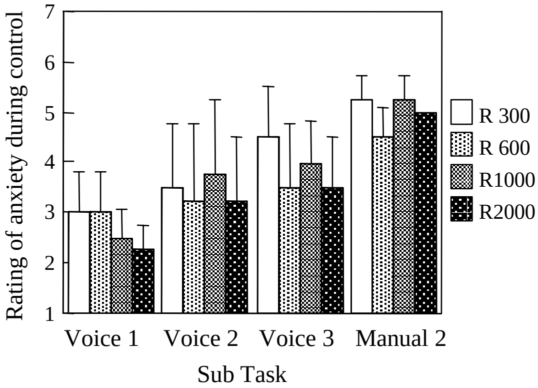


Figure 7. Anxiety during control.

Table 3.
ANOVA (anxiety)

Source	DF	Sum-of-squares	Mean-square	F-Ratio	p-value
Curve(A)	3	3.375	1.125	1.213	0.3150
Sub-task(B)	3	44.875	14.958	16.135	<0.0001
A*B	9	3.250	0.361	0.390	0.9344
Error	48	44.500	0.927		

Difference between means		p-value
Voice 1 - Voice 2	-0.750	0.0324
Voice 1 - Voice 3	-1.188	0.0011
Voice 1 - Manual 2	-2.312	<0.0001
Voice 2 - Voice 3	-0.438	0.2049
Voice 2 - Manual 2	-1.562	<0.0001
Voice 3 - Manual 2	-1.125	0.0018

Mental Workload

Subjects tend to feel the increase of mental workloads as number of voice command increases.

There is no significant difference between mental workload experienced with no task (no voice commands) and two voice commands (Sub-task 1). Therefore, it is assumed that subjects did not feel much load under the condition with two voice commands (sub-task 1).

There is no significant difference in workload experienced between sub-task 2, sub-task 3 and manual sub-task 2. Therefore, it is assumed that subjects feel the load of voice control with more than four commands and the load of four manual operations to be similar.

Subjects tend to feel the increase of mental workloads as curve rate decreases. There is a significant difference in workload experienced between R300 and R2000.

Overall Workloads

An evaluation index of overall workload and NASA-TLX yields similar results proving the effectiveness of NASA-TLX as an evaluation tool.

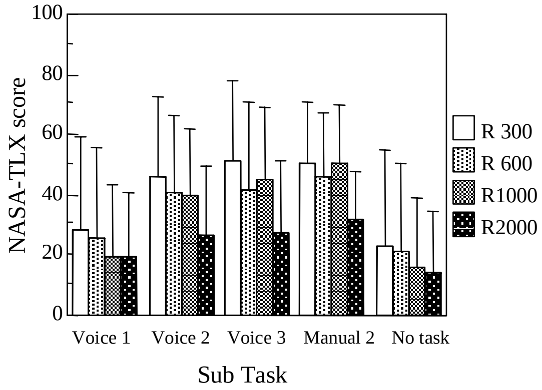


Figure 8. NASA-TLX.

Table 4.
ANOVA (NASA-TLX)

Source	DF	Sum-of-squares	Mean-square	F-Ratio	p-value
Curve(A)	3	2664.68	888.227	1.432	0.2425
Sub-task(B)	4	8650.55	2162.638	3.486	0.0126
A*B	12	728.09	60.675	0.098	>0.9999
Error	60	37227.67	620.461		

Difference between means		p-value
Voice 1 - Voice 2	-15.067	0.0923
Voice 1 - Voice 3	-18.050	0.0448
Voice 1 - Manual 2	-21.442	0.0179
Voice 1 - No sub-task	4.808	0.5871
Voice 2 - Voice 3	-2.983	0.7360
Voice 2 - Manual 2	-6.375	0.4720
Voice 2 - No sub-task	19.875	0.0277
Voice 3 - Manual 2	-3.392	0.7015
Voice 3 -No sub-task	22.858	0.0119
Manual 2 -No sub-task	26.250	0.0041

Difference between means		p-value
R 300 - R 600	4.623	0.5594
R 300 - R1000	5.333	0.5010
R 300 - R2000	15.780	0.0497
R 600 - R1000	0.710	0.9285
R 600 -R2000	11.157	0.1618
R1000 - R2000	10.447	0.1898

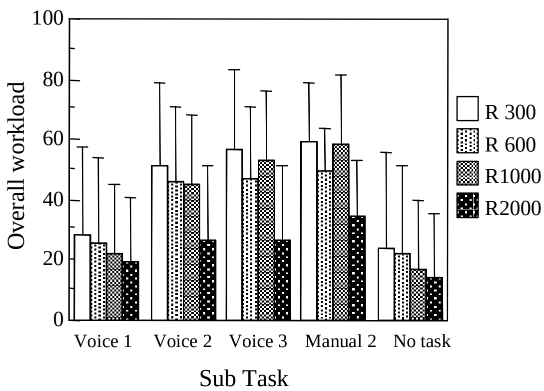


Figure 9. Overall workloads.

Table 5.
ANOVA (overall workloads)

Source	DF	Sum-of-squares	Mean-square	F-Ratio	p-value
Curve(A)	3	4296.85	1432.283	2.418	0.0751
Sub-task(B)	4	12358.17	3089.544	5.215	0.0011
A*B	12	1301.02	108.419	0.183	0.9987
Error	60	35545.50	592.425		

Difference between means		p-value
Voice 1 - Voice 2	-18.750	0.0333
Voice 1 - Voice 3	-22.188	0.0124
Voice 1 - Manual 2	-26.438	0.0032
Voice 1 - No sub-task	4.500	0.6030
Voice 2 - Voice 3	-3.438	0.6910
Voice 2 - Manual 2	-7.688	0.3752
Voice 2 - No sub-task	23.250	0.0090
Voice 3 - Manual 2	-4.250	0.6232
Voice 3 -No sub-task	26.688	0.0029
Manual 2 -No sub-task	30.938	0.0007

Difference between means		p-value
R 300 - R 600	5.950	0.4425
R 300 - R1000	5.000	0.5184
R 300 - R2000	19.750	0.0128
R 600 - R1000	-0.950	0.9022
R 600 -R2000	13.800	0.0780
R1000 - R2000	14.750	0.0601

Vehicle Behavior

Standard Deviation of Steering Angle

There is a correlation between an increase in the number of voice control commands and an increase in the standard deviation of the steering angle. The same correlation exists between an increase in subject anxiety and an increase in the number of voice control commands.

Voice control sub-task 3 indicates an increase in the standard deviation of the steering angle and increased steering instability in comparison with no sub-task case. This fact indicates that even the case of voice control, if the command number increases, the results will have negative influence.

Steering control can be proved to be stable with voice control by comparing manual sub-task 2 and voice sub-task 2. This result supports subjective evaluation.

An influence in the curve rate can be seen between R300 and R2000. A small curve rate increase widens the standard deviation of the steering angle, and also leads to steering control instability. Comparison to other evaluation index, standard deviation of steering angle reflects the influence of curve rate.

Range of Lateral Position

An increase in the number of voice-controlled commands leads to an increase in the range of lateral position, which causes poor driving performance. This is the same phenomenon found with the increases in anxiety, mental workloads, and the standard deviation of the steering angle.

This finding indicates the number of voice-controlled commands has a significant influence on vehicle behavior and also shows drivers need to be careful to limit voice-controlled commands while driving.

The range of lateral position in sub-task 3 is larger in comparison to no sub-task. This data supports the conclusion that even when commands are limited to voice-controlled, if the command number increases lateral positioning is influenced.

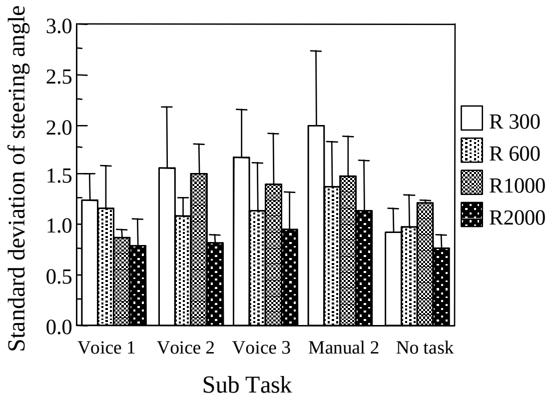


Figure 10. Standard deviation of steering angle.

Table 6.
ANOVA (S.D. of steering angle)

Source	DF	Sum-of-squares	Mean-square	F-Ratio	p-value
Curve(A)	3	3.622	1.207	8.082	0.0001
Sub-task(B)	4	2.990	0.748	5.004	0.0015
A*B	12	1.578	0.131	0.880	0.5712
Error	60	8.964	0.149		

Difference between means		p-value
Voice 1 - Voice 2	-0.229	0.0991
Voice 1 - Voice 3	-0.273	0.0502
Voice 1 - Manual 2	-0.483	0.0008
Voice 1 - No sub-task	0.048	0.7266
Voice 2 - Voice 3	-0.044	0.7478
Voice 2 - Manual 2	-0.254	0.0681
Voice 2 - No sub-task	0.277	0.0472
Voice 3 - Manual 2	-0.210	0.1301
Voice 3 -No sub-task	0.321	0.0221
Manual 2 -No sub-task	0.531	0.0003

Difference between means		p-value
R 300 - R 600	0.332	0.0085
R 300 - R1000	0.192	0.1216
R 300 - R2000	0.584	<0.0001
R 600 - R1000	-0.141	0.2548
R 600 -R2000	0.251	0.0442
R1000 - R2000	0.392	0.0022

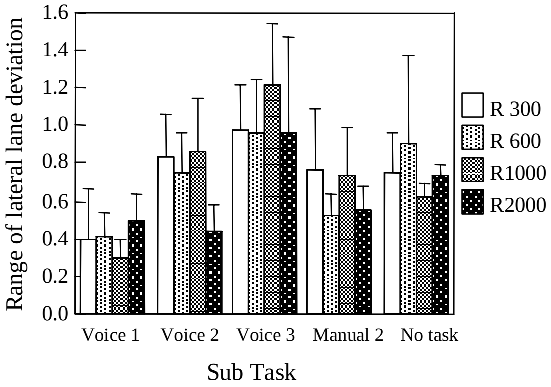


Figure 11. Range of lateral position.

Table 7.
ANOVA (Range of lateral position)

Source	DF	Sum-of-squares	Mean-square	F-Ratio	p-value
Curve(A)	3	0.128	0.043	0.655	0.5837
Sub-task(B)	4	2.822	0.706	10.790	<0.0001
A*B	12	0.742	0.062	0.946	0.5104
Error	50	3.269	0.065		

Difference between means		p-value
Voice 1 - Voice 2	-0.335	0.0011
Voice 1 - Voice 3	-0.623	<0.0001
Voice 1 - Manual 2	-0.248	0.0134
Voice 1 - No sub-task	-0.366	0.0004
Voice 2 - Voice 3	-0.288	0.0044
Voice 2 - Manual 2	0.087	0.3738
Voice 2 - No sub-task	-0.031	0.7469
Voice 3 - Manual 2	0.375	0.0003
Voice 3 -No sub-task	0.257	0.0105
Manual 2 -No sub-task	-0.118	0.2274

CONCLUSIONS

The testing and analysis conclude that increased numbers of voice-controlled commands lead to increased anxiety and mental workloads. Additionally, the analysis also indicates an increased number of voice-controlled commands leads to a negative influence on vehicle behavior.

Although anxiety workloads remain small when operating a higher number of voice-controlled commands, vehicle behavior is negatively influenced same as the manual operation. When comparing manual control and voice control, visual distraction decreases with voice control, but overall distraction does not decrease much.

The results conclude that mental workloads increase when curve rate is small. Additionally, the test proves steering control to be unstable when curve rate is small in standard deviation of steering angle of vehicle behavior index.

REFERENCES

[1] "Japan Automobile Manufactures Association Guideline for In-vehicle Display System - Version 2.1", 2000.

[2] Haga, S., & Mizukami, N., 1996. "Japanese version of NASA Task Load Index: Sensitivity of its workload score to difficulty of three different laboratory tasks", The Japanese Journal of Ergonomics, Vol.32, No.2, 71-79. (In Japanese with English abstract).